

A Project on

Vandermonde Matrices

Submitted for partial fulfillment of award of

MASTER OF SCIENCE

University of Mumbai

Degree in

MATHEMATICS

By

Archana Singh

Sagarika Parmar

Under the Guidance of

Dr. Selby Jose

Department of Mathematics

The Institute of Science

Mumbai - 400032

December, 2022

Declaration

I hereby declare that the project work entitled "**Vandermonde Matrices**" carried out at the Department of Mathematics, The Institute of Science, Mumbai, is a record of an original work done by me under the guidance of **Dr. Selby Jose**, The Institute of Science, and this project work is submitted in the partial fulfillment of the requirements for the award of the degree of Master of Science in Mathematics, University of Mumbai. The results embodied in this report have not been submitted to any other University or Institute for the award of any degree or diploma.

Archana Singh

(Candidate)

Seat No. 202105

Sagarika Parmar

(Candidate)

Seat No. 202131

Certificate

This is to certify that the project report entitled **Vandermonde Matrices**, carried out at the Department of Mathematics, The Institute of Science, Mumbai, in partial fulfillment for the award of the degree of Master of Science in Mathematics, University of Mumbai, is a record of bonafide work carried out by **Ms. Archana Singh** and **Ms. Sagarika Parmar**, Seat No. **202105** and **202131**, under the supervision and guidance of **Dr. Selby Jose** during the academic year 2022 – 2023.

Dr. Selby Jose
Project Guide

External Examiner

Head of Dept.
Mathematics

Place: Mumbai

Date:

Abstract

Matrices whose rows (or columns) consist of monomials of sequential powers are called Vandermonde matrices and can be used to describe several useful concepts and have properties that can be helpful for solving many kinds of problems. In this paper, we will discuss this matrix and some of its properties. This paper investigates the transient solution of discrete and continuous time Markov chains (DTMC and CTMC) using the class of methods based on the undetermined coefficients approach. Two methods that belong to this class are investigated: The well known Eigenvectors method, and what it is referred in the paper as the Vandermonde method, because it leads to the solution of a Vandermonde system of equations. Even if the Vandermonde method is possibly the simplest method to obtain the transient solution of a Markov chain, it has received few attention in the literature. This paper fills this gap, showing its advantages. The Vandermonde method is also exploited to derive interesting relations between a CTMC and its uniformized chain. Based on them, it is proposed a simple, yet powerful method to compute the transient solution of Markov chains referred to as the Uniformized Vandermonde method.

Contents

1 Introduction

2 Basic Definitions

2.1 The Vandermonde Determinant

2.2 The Inverse of Vandermonde Matrix

3 Transient Solution of Markov Chains using the Uniformized Vandermonde Method

3.1 Introduction

3.2 Discrete Time Markov Chains

3.2.1 Undetermined Coefficients of a DTMC

3.2.2 Vandermonde Method

3.3 Continuous Time Markov Chains

3.3.1 Undetermined Coefficients of a CTMC

3.3.2 Vandermonde Method

4 Uniformized Vandermonde Method

5 References

1 Introduction

Probabilistic modeling using Markov chains has been successfully exploited in almost all fields of modern applied mathematics. The stationary solution of a Markov chain is more easy to compute than the transient solution, and it is enough in many cases. However, some applications, as reliability modeling, are primarily interested in the transient solution. In the literature there have been proposed many methods to compute the transient solution of Markov chains. Some examples are the approaches based on Laplace transform techniques, the exponential matrix, differential equation solvers, etc. However, most of these algorithms are not general, are difficult to implement, or are applicable only to chains with a small number of states. We will study the methods based on the undetermined coefficients approach. This approach consists on making an intelligent guess of a function with constant coefficients for the solution of an equation, and then use boundary conditions to solve these coefficients. Another possibility is applying the undetermined coefficients approach directly to the solution of the Markov chain, expressed in terms of its eigenvalues. By doing so, the undetermined coefficients are obtained by solving a confluent Vandermonde system of equations. Furthermore, in this paper it is proposed a novel method that combines the Vandermonde method with the well known technique of Uniformization. Thorough the paper, this novel proposal will be referred to as the Uniformized Vandermonde method. The paper derives the equations required to use the Vandermonde method for both DTMC and CTMC. Simple examples are given that show the simplicity of the Vandermonde method to solve small systems without the help of a computer.

2 BASIC DEFINITIONS

The Vandermonde matrix is a well-known matrix with a very special form that appears in many different circumstances, a few examples are polynomial interpolation least square regression optimal experiment design construction of error-detecting and error-correcting codes as well as more recent work such as determining if a market with a finite set of traded assets is complete calculation of the discrete Fourier transform and related transforms such as the fractional discrete Fourier transform the quantum Fourier transform and the Vandermonde transform solving systems of differential equations with constant coefficients various problems in mathematical- nuclear and quantum physics and describing properties of the Fisher information matrix of stationary stochastic processes .The Vandermonde matrix The Vandermonde matrix is a well-known matrix with a very special form that appears in many different circumstances, a few examples are polynomial interpolation least square regression optimal experiment design construction of error-detecting and error-correcting codes as well as more recent work such as determining if a market with a finite set of traded assets is complete calculation of the discrete Fourier transform and related transforms such as the fractional discrete Fourier transform the quantum Fourier transform and the Vandermonde transform solving systems of differential equations with constant coefficients various prob-

lems in mathematical nuclear and quantum physics and describing properties of the Fisher information matrix of stationary stochastic processes It is known that the Vandermonde matrix is defined by

$$V = V(c_1, \dots, c_m) = \begin{pmatrix} 1 & c_1 & c_1^2 & \dots & c_1^{m-1} \\ 1 & c_2 & c_2^2 & \dots & c_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & c_m & c_m^2 & \dots & c_m^{m-1} \end{pmatrix}$$

2.1 The Vandermonde Determinant

The Vandermonde determinant Often it is not the Vandermonde matrix itself that is useful, instead it is the multivariate polynomial given by its determinant that is examined and used. The determinant of the Vandermonde matrix is usually called the Vandermonde determinant (or Vandermonde polynomial or Vandermondian [168]) and can be written using an exceptionally simple formula. But before we discuss the Vandermonde determinant we will discuss the general determinant.

Definition The determinant is a function of square matrices over a field F to the field F ,

$$\det : M_{n*n}(F) \rightarrow F$$

such that if we consider the determinant as a function of the columns

$$\det(M) = \det(M_1, M_2, \dots, M_n)$$

of the matrix the determinant must have the following

- The determinant must be multi linear $\det(M_1, \dots, aM_k + bN_k, \dots, M_n) = a \det(M_1, \dots, M_k, \dots, M_n) + b \det(M_1, \dots, N_k, \dots, M_n)$.
- The determinant must be alternating, that is if

$$M_i = M_j$$

for some $i = j$ then $\det(M) = 0$

- If I is the identity matrix then $\det(I) = 1$

2.2 The Inverse of Vandermonde Matrix

The inverse for the Vandermonde matrix has been known for a long time, especially since the solution to a Lagrange interpolation problems gives the inverse indirectly. Here we will only give a short overview of the work on expressing the inverse as an explicit matrix.

An explicit expression for the inverse matrix has been known since at least the end of the 1950s.

Theorem: The elements of the inverse of an n-dimensional Vandermonde matrix V can be calculated by:

$$(V_N^{-1})_{ij} = \frac{(-1)^{j-1} a_{n-j,i}}{\prod_{k=1}^n (x_k - x_i)}$$

where $a_{j,i}$ is the j:th elementary symmetric polynomial with variable x_i set to zero.

3 Transient Solution of Markov Chains Using the Uniformized Vandermonde Method

3.1 Introduction

Probabilistic modeling using Markov chains has been successfully exploited in almost all fields of modern applied mathematics. The stationary solution of a Markov chain is more easy to compute than the transient solution, and it is enough in many cases. However, some applications, as reliability modeling, are primarily interested in the transient solution.

We will study the methods based on the undetermined coefficients approach. This approach consists on making an intelligent guess of a function with constant coefficients for the solution of an equation, and then use boundary conditions to solve these coefficients. The undetermined coefficients approach has been successfully used to find the solution of many difference and differential equations. The transient solution of a Discrete or Continuous Time Markov Chain, DTMC and CTMC, is the solution of a difference or differential equation respectively. Therefore, DTMC and CTMC are suited to be solved using algorithms based on the undetermined coefficients approach.

3.2 Discrete Time Markov Chains

Definition: A Discrete Time Markov chain is a mathematical process that transitions from one state to another within a finite number of possible states. It is a collection of different states and probabilities of a variable, where its future condition or state is substantially dependent on its immediate previous state. for every $i \in S, j \in S$,

$$P\{X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}\} = P\{X_{n+1} = j | X_n = i\}$$

and the process is called a DTMC.

Besides the Markov property there is another property that would need to be satisfied in order to be able to effectively analyze a DTMC. A DTMC is called time-homogeneous if

the probability of transitioning from a state to another state does not vary with time. In other words, for a DTMC with state space S to be time-homogeneous, for every $i \in S, j \in S$,

$$P\{X_{n+1} = j | X_n = i\} = P\{X_1 = j | X_0 = i\}.$$

We denote the above expression as p_{ij} the one step transition probability of going from state i to j . Using the transition probabilities p_{ij} for every $i \in S$ and $j \in S$, we can build a square matrix

$$P = [p_{ij}].$$

The matrix P is called transition probability matrix. The sum of the elements of each row adds to 1 because given that the DTMC is in state i the next state ought to be one of the states in S hence for every $i \in S$,

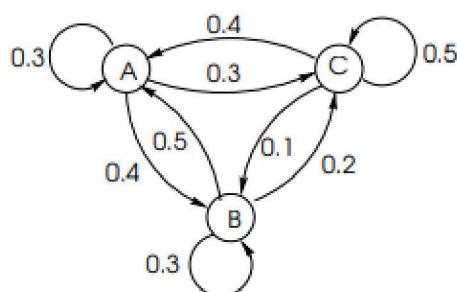
$$\sum_{j \in S} P_{ij} = 1$$

Now to understand it better, let us consider few examples.

Example 1: Consider three long-distance telephone companies A, B and C. Everytime a sale is announced, users switch from one company to another. Let X_n denote the long-distance company with which a particular user named Jill is just before the n th sale announcement. We have $S = \{A, B, C\}$. Based on the customers' switching in the past it is estimated that Jill's switching patterns follow the following transition probability matrix:

$$P = \begin{matrix} & \begin{matrix} A & B & C \end{matrix} \\ \begin{matrix} A \\ B \\ C \end{matrix} & \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.5 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{bmatrix} \end{matrix}$$

For example, if Jill is with B before a sale is announced, she would switch to A with probability 0.5 and C with probability 0.2 or stay with B with probability 0.3 as evident from the second row in P . Now, converting the above P matrix into a transition diagram, we get the picture shown below.



We now know how to form a transition matrix for a given DTMC problem. There are various methods to solve the system of equation but our focus is to solve this system using the Vandermonde Method.

Theorem 2.2.1: Let $\lambda_1, l = 1 \dots L$ be the eigenvalues of P, each with multiplicity k_l . $\sum_l k_l = k$

Without loss of generality, assume a possible eigenvalue $1 = 0$ with multiplicity k_1 . Then:

$$\pi_j(n) = \sum_{m=0}^{k_1-1} a_j^{(l,m)} \delta_{n-m} + \sum_{l=2}^L \lambda_l^n \sum_{m=0}^{k_l-1} a_j^{(l,m)} n^m \longrightarrow (1)$$

where δ_k is the Kronecker's delta ($\delta_k = 1$ for $k = 0$ and 0 otherwise).

Before we prove the theorem, let us go through preliminaries to understand the proof better.

PRELIMINARIES

i) Jordan Block

A $d \times d$ matrix $J_{d,\lambda}$ is said to be a Jordan block of dim d and eigenvalue λ iff its diagonal entries are all equal to λ , its super diagonal entries are all equal to 1, and all its other entries are equal to 0.

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{bmatrix}.$$

Note: A Jordan block is upper triangular, and the diagonal entries of an upper triangular matrix are equal to its eigenvalues. This is the reason why λ is called the eigenvalue of the Jordan block $J_{d,\lambda}$.

ii) Jordan Form A square matrix is said to be in Jordan form if it is block diagonal where each block is a Jordan block.

$$J_x = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Proof:

Using the Jordan Form of the matrix $P = AJA^{-1}$, where A is some invertible matrix, and $J = \text{diag}(J_1, J_2, \dots, J_L)$ is a diagonal block matrix, with the Jordan blocks J_l (J_l has λ_l in the main diagonal, 1s in the first upper diagonal, and size equal to the multiplicity of λ_l, k_l):

$$J_l^{k_l \times k_l} = \begin{bmatrix} \lambda_l & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_l & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_l & 1 & \cdots & 0 & 0 \\ \cdots & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & \lambda_l & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_l \end{bmatrix} \longrightarrow (2)$$

We have that $J^n = \text{diag}(J_1^n, J_2^n, \dots, J_L^n)$. To compute the powers of a Jordan block, we write:

$$J_l = \lambda_l I + U_1^{k_l * k_l} \longrightarrow (3)$$

where $U_1^{k_l * k_l}$ is a matrix with the first upper diagonal equal to 1, and the other elements equal to zero:

$$U_1^{k_l \times k_l} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \longrightarrow (4)$$

It can be easily obtained that :

$$(U_1^{k_l * k_l})^n = U_n^{k_l * k_l} \longrightarrow (5)$$

where $U_n^{k_l * k_l}$ is a matrix with the n-th upper diagonal equal to 1, and the other elements equal to zero. $U_n^{k_l * k_l} = 0$ if $n \geq k_l$.

Assume first that $\lambda_l \neq 0, \forall l$. We can write:

$$J_l^n = (\lambda_l I + U_1^{k_l * k_l})^n = \sum_{i=0}^n \binom{n}{i} \lambda_l^{n-i} (U_1^{k_l * k_l})^i = \lambda_l^n * \begin{bmatrix} 1 & \lambda_l^{-1} \binom{n}{1} & \lambda_l^{-2} \binom{n}{2} & \cdots & \lambda_l^{-(k_l-1)} \binom{n}{k_l-1} \\ 0 & 1 & \lambda_l^{-1} \binom{n}{1} & \cdots & \lambda_l^{-(k_l-2)} \binom{n}{k_l-2} \\ 0 & 0 & 1 & \cdots & \lambda_l^{-(k_l-3)} \binom{n}{k_l-3} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \longrightarrow (6)$$

where it is assumed $\binom{k}{j} = 0$ for $j \geq k$, and $(U_1^{k_l * k_l})^0 = I$.

$$\text{Since } \binom{n}{k_l-1} = \frac{n(n-1)\dots(n-k_l)+2}{(k_l-1)!} = \frac{1}{(k_l-1)!} (n^{k_l-1} + \dots)$$

is a polynomial in n of degree $k_l - 1$, we conclude that the elements of the matrix $P^n = AJ^nA^{-1}$ will be linear combinations of the eigenvalues of P to the power of n, times polynomial in n of degree equal, at most, to their multiplicity = 1, and so it will be $\pi(0)P^n$, as stated in equation (1).

3.2.1 Undetermined Coefficients of a DTMC

In the equation (1) there are up to k determined coefficients $a_j^{(l,m)}$ of $\pi_j(n)$ to be determined. we will focus on computation of these coefficients. The sentence UC of $\pi_j(n)$ or simply UC will be used to refer the undetermined coefficients. Notation to denote the UC of $\pi_j(n)$ is:

$$u_j = [u_j^{(1)} \dots u_j^{(L)}] \longrightarrow (A)$$

where u_j is a column vector with $u_j^{(l)} = [a_j^{(l,0)} \dots a_j^{(l,k_l-1)}]$, $l=1\dots L$. In the following two methods to compute the UC of $\pi_j(n)$ are described.

3.2.2 Vandermonde Method

The undetermined coefficients of $\pi_j(n)$ can be obtained solving the system of equations that results from imposing the boundary conditions to equation (1).

$$\pi(n) = (\pi(0)P^n)_j \longrightarrow (7)$$

for $n = 0, 1, \dots$, up to the number of UC to be determined. If we don't exploit the fact that the geometric multiplicity of some eigenvalues may be larger than 1, then there will be k UC, where k is the number of states of the Markov chain, i.e. the size of the square matrix P .

We can obtain vector u_j with UC of $\pi_j(n)$ by solving the below system of equation:

$$[\mathbf{A}_1 \ \dots \ \mathbf{A}_L] \mathbf{u}_j = \mathbf{B} \longrightarrow (8)$$

where the sub-matrices $A_l^{k \times k_l}$, $l=1\dots L$, ($\lambda_1 = 0, \lambda_l \neq 0$) are given by:

$$\mathbf{A}_1^{k \times k_1} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & & & \end{bmatrix} \longrightarrow (9)$$

$$\mathbf{A}_l^{k \times k_l} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_l & \lambda_l & \dots & \lambda_l \\ \lambda_l^2 & 2\lambda_l^2 & \dots & 2^{k_l-1}\lambda_l^2 \\ \dots & & & \\ \lambda_l^n & n\lambda_l^n & \dots & n^{k_l-1}\lambda_l^n \\ \dots & & & \end{bmatrix} \longrightarrow (10)$$

and B is the column vector

$$\mathbf{B} = [\pi_j(0) \ (\pi(0)\mathbf{P})_j \ (\pi(0)\mathbf{P}^2)_j \dots] \longrightarrow (11)$$

Note that if the chain starts in state i with probability 1 (i.e., (0) is a probability vector with the probability 1 in the component i), then $B = [\delta_{ij}(P)_{ij}(P^2)_{ij} \dots]$ where δ_{ij} is the Kronecker's delta ($\delta_{ij}= 1$ for $i = j$, and 0 otherwise) .

Note: The Vandermonde method allows obtaining the solution of small chains without the help of a computer and it also gives a more general solution.

Example 1 Find the probability of being in state 3 in n steps starting from state 1 (i.e. $\pi_3(n)$ with $\pi(0) = [1 \ 0 \ 0 \ 0]$). Assume the DTMC with

$$\mathbf{P} = \begin{bmatrix} 4/5 & 1/5 & 0 \\ 0 & 4/5 & 1/5 \\ 1/5 & 1/5 & 3/5 \end{bmatrix}$$

Solution: We first find the eigenvalues of P

$$|\mathbf{P} - I\lambda| = \begin{vmatrix} \frac{4}{5} - \lambda & \frac{1}{5} & 0 \\ 0 & \frac{4}{5} - \lambda & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{3}{5} - \lambda \end{vmatrix} = 0$$

$$\Rightarrow -1/25(\lambda - 1)(5\lambda - 3)^2 = 0$$

$$\lambda = 1, 3/5, 3/5 \text{ i.e. } \lambda_1=1, \lambda_2=3/5$$

We see that multiplicity of 1 is one and $3/5$ is two Following eq (1) , we get

$$\pi_3(n) = \lambda_1^n A + \lambda_2^n (B + Cn)$$

Note: Here Kronecker's delta is 1 as multiplicity of $\lambda_1 = 0$

$$= A + (3/5)^n (B + Cn) \longrightarrow (12)$$

imposing the boundary conditions , we get

$$\pi_3(0) = A + B = (P^0)_{13} = 0$$

$$\pi_3(1) = A + (3/5)^1 (B + C) = (P^1)_{13} = 0$$

$$\pi_3(2) = A + (3/5)^2 (B + C) = (P^2)_{13} = 1/25$$

On solving the above equations, we get

$$A = 1/4, B = 1/4, C = -1/6$$

Substituting these values in eq (7)

$$\pi_3(n) = 1/4 - (3/5)^n (1/4 + n/6) \text{ for } n \geq 0$$

Example 2 Find the probability of being in state 3 in n steps starting from state 1 starting from state 1 (i.e. $\pi_3(n)$ with $\pi(0) = [1 \ 0 \ 0 \ 0 \ 0]$). Assume the DTMC with

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 1/3 & 2/3 \end{bmatrix}$$

Solution: We first find the eigenvalues of P

$$|P - I\lambda| = \begin{vmatrix} -\lambda & 1 & 0 & 0 \\ 0 & -\lambda & 1 & 0 \\ 0 & 0 & \frac{2}{3} - \lambda & 1/3 \\ 0 & 0 & 1/3 & \frac{2}{3} - \lambda \end{vmatrix} = 0$$

$$\Rightarrow 1/3(\lambda - 1)\lambda^2(3\lambda - 1) = 0$$

$$\lambda = 0, 0, 1, 1/3$$

$$\text{i.e. } \lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 1/3$$

Note: Here the multiplicity of $\lambda_1=2$ i.e. Kronecker's delta $\delta_2 = 1$

$$\pi_3(n) = \delta_n A + \delta_{n-1} B + C + D(1/3)^n \longrightarrow (13)$$

Imposing the boundary conditions, We get

$$\pi_3(0) = A + C + D = (P^0)_{13} = 0$$

$$\pi_3(1) = B + C + D/3 = (P^1)_{13} = 0$$

$$\pi_3(2) = C + D/3^2 = (P^2)_{13} = 1$$

$$\pi_3(3) = C + D/3^3 = (P^3)_{13} = 2/3$$

Solving the above equations , We get

$$A = -5, B = -2, C = 1/2, D = 9/2$$

Substituting these values in eq (8)

$$\pi_3(n) = -5 \delta_n - 2 \delta_{n-1} + 1/2 + 9/2 (1/3)^n \text{ for } n \geq 0$$

3.3 Continuous Time Markov Chains

A continuous-time Markov chain (CTMC) is a continuous stochastic process in which, for each state, the process will change state according to an exponential random variable and then move to a different state as specified by the probabilities of a stochastic matrix.

A continuous time stochastic process with state space S is said to be a CTMC if it satisfies Markov's property

$$P\{X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}\} = P\{X_{n+1} = j | X_n = i\}$$

for every $i \in S, j \in S$.

Theorem: Let $\lambda_l, l = 1, L$ be the eigenvalues of Q, each with multiplicity $k_l (\sum_l k_l = k)$, then:

$$\pi_j(n) = \sum_{l=1}^L e^{\lambda_l t} \sum_{m=0}^{k_l-1} a_j^{(l,m)} t^m \longrightarrow (14)$$

Before we prove the theorem , lets us go through necessary preliminaries to understand the proof better.

PRELIMINARIES

We know that an operator on a finite-dimensional complex vector space can sometimes be represented by a diagonal matrix i.e. exactly when the operator is diagonalizable. Note that for a diagonalizable operator, the Jordan form itself only consists of 1×1 Jordan blocks and is diagonal; this follows from the fact that for a diagonalizable operator, each generalized eigenvectors is a usual eigenvector. To see an application, recall that for an operator T , we define the operator e^T by

$$e^t = I + \sum_{k=1}^{\infty} \frac{T^k}{k!} \longrightarrow (B)$$

The motivation for this of course comes from the power series expansion of e^x . In general this may not be so easy to actually compute, but we will see the Jordan forms give us a nice way of describing e^T via a matrix. If A is a square matrix, we have the same definition. If A is block diagonal:
If A is a square matrix, we have the same definition. If block diagonal:

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix}$$

it turns out that A^k is also block-diagonal and

$$A^k = \begin{pmatrix} A_1^k & & \\ & \ddots & \\ & & A_m^k \end{pmatrix}.$$

$$e^A = \begin{pmatrix} e^{A_1} & & \\ & \ddots & \\ & & e^{A_m} \end{pmatrix}$$

So, if we can represent T by a block diagonal matrix, we can describe e^T by computing the exponential of each block. We know that if a block is diagonal, computing its exponential is easy. The point is that it is also easy for Jordan blocks! For simplicity, consider the 2×2 Jordan block

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

$$J^2 = \begin{pmatrix} \lambda^2 & 2\lambda \\ 0 & \lambda^2 \end{pmatrix}$$

$$J^3 = \begin{pmatrix} \lambda^3 & 3\lambda^2 \\ 0 & \lambda^3 \end{pmatrix}$$

and in general,

$$J^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}$$

Thus ,

$$\begin{aligned}
e^J &= I + \sum_{k=1}^{\infty} \frac{J^k}{k!} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k=1}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{k=1}^{\infty} \begin{pmatrix} \frac{\lambda^k}{k!} & \frac{\lambda^{k-1}}{(k-1)!} \\ 0 & \frac{\lambda^k}{k!} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} & \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ 0 & \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \end{pmatrix} \\
&= \begin{pmatrix} e^\lambda & e^\lambda \\ 0 & e^\lambda \end{pmatrix}
\end{aligned}$$

Thus we know exactly what the exponential of a 2×2 Jordan block is. Similarly, one can show that for a size k Jordan block, its exponential will be the uppertriangular matrix which as e^λ down the main diagonal, $e^\lambda/1!$ down the diagonal above that, $e^\lambda/2!$ down the diagonal above that, $e^\lambda/3!$ down the diagonal above that, and so on. Thus, we know exactly what the exponential of any Jordan block is, and hence what the exponential of any Jordan form is. So, we can always find a way to explicitly express e^T via a matrix representation.

Proof: We proceed in similar way as in Theorem 2.2.1 using the Jordan form of the matrix $Q = AJA^{-1}$, where A is some invertible matrix, and $J = \text{diag}(J_1, J_2, \dots, J_L)$ is a diagonal block matrix, with Jordan blocks J_l . We have :

$$e^{Qt} = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!} = Ae^{Jt}A^{-1} \longrightarrow (15)$$

where $e^{Jt} = \text{diag}(e^{J_1 t}, e^{J_2 t}, \dots, e^{J_L t})$. To compute the exponential matrices of the Jordan blocks, we write: $J_l = \lambda_l I + U_1^{k_l * k_l}$, where the matrix $U_1^{k_l * k_l}$ is as define in Theorem 2.2.1. Thus:

$$\begin{aligned}
e^{J_l t} &= \sum_{n=0}^{\infty} \frac{(J_l t)^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{n!} (\lambda_l I + U_1^{k_l * k_l})^n \\
&= \sum_{n=0}^{\infty} \frac{(J_l t)^n}{n!} \sum_{i=0}^n \binom{n}{i} (\lambda_l^{n-i} (U_1^{k_l * k_l})^i) \\
&= \sum_{i=0}^{\infty} \sum_{n=i}^{\infty} \frac{(\lambda_l t)^{n-i}}{n!} \binom{n}{i} (U_1^{k_l * k_l t})^i \\
&= \sum_{i=0}^{\infty} \frac{(U_1^{k_l * k_l t})^i}{i!} \sum_{j=0}^{\infty} \frac{(\lambda_l t)^j}{j!} \\
&= e_l^{\lambda_l t} \sum_{i=0}^{k_l-1} \frac{(U_1^{k_l * k_l t})^i}{i!} t^i \longrightarrow (16)
\end{aligned}$$

where it was used the change $j = n - i$, and the fact that $(U_1^{k_l * k_l t})^i = 0$ if $i \geq k_l$. Using (16), we conclude that the elements of the matrix $e^{Qt} = Ae^{Jt}A^{-1}$ will be linear combination of $e^{\lambda_l t}$ times polynomials in t of degree equal , at most, to $k_l - 1$, as stated in equation (14).

3.3.1 Undetermined Coefficients of a CTMC

In equation (14) there are up to k UCs $a_j^{(l,m)}$ of $\pi_j(t)$ to be determined. We will use the same column vector as in DTMC i.e. equation (A).

3.3.2 Vandermonde method

The UC of $\pi_j(t)$ can be obtained solving the system of equations that results from imposing the boundary conditions to equation (14):

$$\frac{\partial^n \pi_j(0)}{\partial t^n} = (\boldsymbol{\pi}(0) \mathbf{Q}^n)_j \longrightarrow (17)$$

for $n=0, 1, 2, \dots$, up to the number of UC to be determined. As in the discrete case, the vector u_j with the UC of $\pi_j(t)$, can be obtained solving the system of equations: $[A_1 \dots A_L] u_j = B$, where the sub matrices $A_l^{k \times k_l}$, $l=1 \dots L$ are given by:

$$\mathbf{A}_l^{k \times k_l} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_l & 1 & \dots & 0 \\ \lambda_l^2 & 2\lambda_l & \dots & 0 \\ \dots & & & \\ \lambda_l^n & n\lambda_l^{n-1} & \dots & n^{k_l-1}\lambda_l^{n-(k_l-1)} \\ \dots & & & \end{bmatrix} \longrightarrow (18)$$

where $n^m = n(n-1)\dots(n-m+1)$, $n^m = 0$ for $n \leq m$, and B is the column vector:

$$\mathbf{B} = [\pi_j(0) \quad (\boldsymbol{\pi}(0) \mathbf{Q})_j \quad (\boldsymbol{\pi}(0) \mathbf{Q}^2)_j \dots] \longrightarrow (19)$$

Note:

- If the chain starts in state i with probability 1 (i.e., (0) is a probability vector with the probability 1 in the component i), then $B = [\delta_{ij}(Q)_{ij}(Q^2)_{ij}]$.
- The vector B given by (19) is exactly the same form than in the DTMC (equation (11)).

Example

Find the probability of being in state 3 at time t starting from state 1 (i.e. $3(t)$ with $\boldsymbol{\pi}(0) = [1 \ 0 \ 0]$). Assume a CTMC with

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

solution: We will first find the eigenvalues of given matrix i.e.

$$|\mathbf{Q} - \lambda I| = \begin{vmatrix} -1 - \lambda & 1 & 0 \\ 0 & -1 - \lambda & 1 \\ 1 & 1 & -2 - \lambda \end{vmatrix}$$

$$\implies -\lambda(\lambda + 2)^2 = 0$$

$$\lambda = 0, -2 \text{ i.e. } \lambda_1 = 0, \lambda_2 = -2$$

we see that multiplicity of 0 is 1 and that of -2 is 2, following equation (14), We get:

$$\boldsymbol{\pi}3(t) = e^{\lambda_1 t} A + e^{\lambda_2 t} (B + Ct) \longrightarrow (20)$$

imposing boundary conditions:

$$\pi_3(0) = A + B = (Q^0)_{13} = 0$$

$$\pi'_3(1) = -2B + C = (Q^1)_{13} = 0$$

$$\pi''_3(2) = (-2)^2 B + 2(-2)C = (Q^2)_{13} = 1$$

We get,

$$A = -B = 1/4, C = -1/2$$

Thus substituting in eq (20), We get,

$$\pi_3(t) = 1/4 + e^{-2}t(1/4 + t/2) \text{ for } t \geq 0$$

4 Uniformized Vandermonde Method

In case of the Vandermonde method, the eigenvalues of the matrix \mathbf{Q} may be out of the unit circle, and thus, the matrix used to compute the undetermined coefficients may have a large norm and be ill-conditioned. This problem can be alleviated using the well known uniformization method . This method consists of considering the uniformized matrix:

$$\mathbf{P} = \mathbf{I} + \frac{1}{q}\mathbf{Q} \quad (21)$$

where q is a constant such that $\max_i |(\mathbf{Q})_{ii}| \leq q < \infty$, and using the equation (conveniently truncated):

$$\pi(t) = \pi(0)e^{\mathbf{Q}t} = \sum_{n=0}^{\infty} e^{-qt} \frac{(qt)^n}{n!} \pi(0)\mathbf{P}^n \quad (22)$$

Note that the matrix \mathbf{P} defined by (21) is stochastic, and the equation (22) can be easily proved by direct substitution of the definition of the matrix \mathbf{P} (21). The problem of directly applying the uniformization formula (22) is that computing the powers \mathbf{P}^n can be costly for large number of states. Additionally, storing these powers will be infeasible for large matrices, thus, the powers \mathbf{P}^n will have to be computed for each time t that we wish to evaluate.

To cope with the above mentioned problems we can use the DTMC defined by \mathbf{P} to compute the uniformized chain probabilities $\pi_j^{(P)}(n)$ for each of the states j we are interested. Thorough this section the indexes $^{(P)}$ and $^{(Q)}$ will be used to distinguish the uniformized chain and the original CTMC chain with infinitesimal generator \mathbf{Q} , respectively. Note that the initial conditions, $\pi(0)$, are the same for both chains. So, $^{(P)}(Q)$ will not be used with $\pi(0)$. From equation (22) we have:

$$\pi_j^{(Q)}(t) = \sum_{n=0}^{\infty} e^{-qt} \frac{(qt)^n}{n!} \pi_j^{(P)}(n) \quad (23)$$

Let $\lambda_l^{(P)}, l = 1, \dots, L$ be the eigenvalues of the uniformized matrix \mathbf{P} given by equation (21), each with multiplicity k_l . Without loss of generality, let $\lambda_1^{(P)}$ be the eigenvalue equal to 1 , and $\lambda_2^{(P)}$ a possible eigenvalue equal to 0 . Let $b_j^{(l,m)}$ be the UC of $\pi_j^{(P)}(n)$. On substituting into (23) we have:

$$\begin{aligned} \pi_j^{(Q)}(t) &= \sum_{n=0}^{\infty} e^{-Q} \frac{(qt)^n}{n!} \left[b_j^{(1,0)} + \sum_{m=0}^{k_2-1} b_j^{(2,m)} \delta_{n-m} + \sum_{l=3}^L \left(\lambda_i^{(P)} \right)^n \sum_{m=0}^{k_1-1} b_j^{(l,m)} n^m \right] \\ &= b_j^{(1,0)} + e^{-qt} \left[b_j^{(2,0)} + \sum_{m=1}^{k_2-1} \frac{(qt)^m}{m!} b_j^{(2,m)} \right] + \\ &\quad \sum_{l=3}^L e^{-qt(1-\lambda_l^{(P)})} \left[b_j^{(l,0)} + \sum_{m=1}^{k_1-1} b_j^{(l,m)} q^{(l,m)}(t) \right] \end{aligned} \quad (24)$$

where $q^{(l,m)}(t)$ are polynomials in t of degree m :

$$q^{(l,m)}(t) = \sum_{i=1}^m q_i^{(m)} \left(q \lambda_i^{(P)} \right)^i t^i \quad (25)$$

with the coefficients $q_i^{(m)}$ given by the recurrence relation:

$$\begin{aligned} q_i^{(m)} &= \{ 1, i = 1 \\ &= \sum_{k=i-1}^{m-1} \binom{m-1}{k} q_{i-1}^{(k)}, i = 2, \dots, m \end{aligned} \quad (26)$$

Let $\lambda_l^{(Q)}, l = 1, \dots, L$ be the eigenvalues of the infinitesimal generator \mathbf{Q} , and $a_j^{(l,m)}$ the UC of $\pi_j^{(Q)}(t)$. Comparing equation (22) and (32), it turns out that it must be:

$$\lambda_l^{(Q)} = -q \left(1 - \lambda_l^{(P)}\right) \quad (27)$$

$$\lambda_l^{(P)} = 1 + \frac{\lambda_l^{(Q)}}{q} \quad (28)$$

and $a_j^{(l,0)} = b_j^{(l,0)}, l = 1, \dots, L$. Thus, if there are not confluent eigenvalues, the UC of $\pi_j^{(P)}(n)$ and $\pi_j^{(Q)}(t)$ are the same (note that thorough the paper it is used confluent and multiple eigenvalues interchangeably). In the confluent case, from equations (32) and (33) we have that:

$$a_j^{(2,m)} = \frac{q^m}{m!} b_j^{(2,m)}, m = 1, \dots, k_2 - 1 \quad (29)$$

$$a_j^{(l,m)} = \left(q\lambda_l^{(P)}\right)^m \sum_{k=m}^{k_1-1} q_m^{(k)} b_j^{(l,k)}, m = 1, \dots, k_1 \quad (30)$$

Using the definition of eigenvector it is easy to prove that equations (27) and (28) imply that the eigenvectors of \mathbf{Q} and its uniformized matrix \mathbf{P} must be the same.

Summing up, the Uniformized Vandermonde Method proposed in this paper consists of the following steps:

1. Compute the eigenvalues of $\mathbf{Q}, \lambda_l^{(Q)}$.
2. Process the eigenvalues $\lambda_l^{(Q)} < 0$, merging those that are near confluent, if any, and limiting the maximum multiplicity. The multiplicity of the eigenvalue $\lambda_1^{(Q)} = 0, k_1$, should be determined by the number of irreducible closed sets.
3. Choose an appropriate value for the uniformization parameter q .
4. Compute the uniformized matrix $\mathbf{P} = \mathbf{I} + \frac{1}{q}\mathbf{Q}$. To save memory space, the matrix \mathbf{P} can overwrite matrix \mathbf{Q} in this step, since Q is not needed anymore.
5. Using \mathbf{P} and $\pi(0)$, compute the matrix \mathbf{B} using the algorithm. Then, the matrix \mathbf{P} can be removed, since it is not needed anymore.
6. Compute the eigenvalues of $\mathbf{P}, \lambda_l^{(P)}$, using equation.
7. Using $\lambda_l^{(P)}$, construct the Vandermonde matrix.
8. Solve the resulting Vandermonde system to obtain the UC of $\pi_j^{(P)}(n)$. If there aren't confluent eigenvalues, these are the UC of $\pi_j^{(Q)}(t)$, and we are done. In case of confluent eigenvalues, use equations (29) and (30) to compute the remaining UC of $\pi_j^{(Q)}(t)$.
9. Use UC of $\pi_j^{(Q)}(t)$ and the eigenvalues $\lambda_l^{(Q)}$ in equation (22) to evaluate the desired transient solution.

Remark: It is interesting to note that if the uniformized chain probabilities $\pi_j^{(P)}(n)$ given by the matrix \mathbf{P} have a limiting distribution, they converge to $\pi_j^{(Q)}(t)$ at the points $n = [qt]$, where $[x]$ stands for the integral value of x . This comes from the fact:

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \lim_{n \rightarrow \infty} \left(\mathbf{I} + \frac{1}{q}\mathbf{Q}\right)^n = \lim_{n \rightarrow \infty} \left(\mathbf{I} + \frac{t\mathbf{Q}}{n}\right)^n = e^{\mathbf{Q}t} \quad (31)$$

The following example illustrates the Uniformized Vandermonde method, and shows a uniformized chain having no limiting distribution.

Example: Assume a CTMC with

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

We want the probability of being in state $\pi_j^{(Q)}(t), j = 1, 2$ starting from state 1 1 ($\pi(0) = [1 \ 0]$), using the Uniformized Vandermonde method. The eigenvalues of \mathbf{Q} are $\lambda_1^{(Q)} = 0, \lambda_2^{(Q)} = -2$. Choosing the uniformization parameter $q = 1$ we have, $\lambda_1^{(P)} = 1$ and $\lambda_2^{(P)} = -1$, and a uniformized matrix:

$$\mathbf{P} = \mathbf{I} + \frac{1}{q} \mathbf{Q} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

which is clearly periodic. The Vandermonde matrix of the uniformized chain is:

$$\mathbf{V}^{(P)} = \begin{bmatrix} 1 & 1 \\ \lambda_1^{(P)} & \lambda_2^{(P)} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and the matrix \mathbf{B} :

$$\mathbf{B} = \begin{bmatrix} \pi_1(0) & \pi_2(0) \\ (\pi(0)\mathbf{P})_1 & (\pi(0)\mathbf{P})_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, the UC of the uniformized chain are given by:

$$[u_1 \ u_2] = (\mathbf{V}^{(P)})^{-1} \mathbf{B} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

Since the eigenvalues are not confluent, the UC of $\pi_j^{(P)}(n)$ and $\pi_j^{(Q)}(t)$ are the same, so we have:

$$\begin{cases} \pi_1^{(P)}(n) = 1/2 + 1/2(-1)^n \\ \pi_2^{(P)}(n) = 1/2 - 1/2(-1)^n \end{cases} \quad (32)$$

$$\begin{cases} \pi_1^{(Q)}(t) = 1/2 + 1/2e^{-2t} \\ \pi_2^{(Q)}(t) = 1/2 - 1/2e^{-2t} \end{cases} \quad (33)$$

In this case the uniformized chain does not have a limiting distribution, and the equation (31) does not apply: Clearly, equations (32) do not converge to equations (33) at the points $n = [qt]$.

References

- [1] A. Bjorck and V. Pareyra. "Solution of Vandermonde Systems of Equations." , 1970.
- [2] Llorenc Cerdà - Alabern, University Politecnica de Catalunya Computer Architecture Dept, 2011.
- [3] Angela Peace , biomathematics, MATH 5355 Spring , 2017.
- [4] Renato Lo Cigno , Simulation and Performance , 2018-19.